

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Publications, Agencies and Staff of the U.S.
Department of Commerce

U.S. Department of Commerce

2012

Validation of the GOES-R ABI flood and standing water algorithm using gauging station measurements and interpretation maps

Rui Zhang

George Mason University, rzhang5@gmu.edu

Donglian Sun

George Mason University

Yunyue Yu

NOAA/NESDIS

Anthony Stefanidis

George Mason University

Mitchell D. Goldberg

NOAA/NESDIS

Follow this and additional works at: <https://digitalcommons.unl.edu/usdeptcommercepub>



Part of the [Environmental Sciences Commons](#)

Zhang, Rui; Sun, Donglian; Yu, Yunyue; Stefanidis, Anthony; and Goldberg, Mitchell D., "Validation of the GOES-R ABI flood and standing water algorithm using gauging station measurements and interpretation maps" (2012). *Publications, Agencies and Staff of the U.S. Department of Commerce*. 364.
<https://digitalcommons.unl.edu/usdeptcommercepub/364>

This Article is brought to you for free and open access by the U.S. Department of Commerce at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications, Agencies and Staff of the U.S. Department of Commerce by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Validation of the GOES-R ABI flood and standing water algorithm using gauging station measurements and interpretation maps

Rui Zhang ^{a,*}, Donglian Sun ^a, Yunyue Yu ^b, Anthony Stefanidis ^a, Mitchell D. Goldberg ^c

^a Department of Geography and Geoinformation Science, College of Science, George Mason University, Fairfax, VA, 22030, USA

^b NOAA/NESDIS Center for Satellite Application and Research, College Park, MD 20742, USA

^c NOAA/NESDIS Center for Satellite Application and Research, Camp Springs, MD 20746, USA

ARTICLE INFO

Article history:

Received 10 December 2011

Received in revised form 10 April 2012

Accepted 14 April 2012

Available online xxxx

Keywords:

validation

flood detection

gauging station data

human interpretation

MODIS

ABSTRACT

Validation is an important task in the development of satellite remote sensing products. Strategies for validation vary depending on the nature of the products. The validation process of the flood and standing water product (FSW) for the Geostationary Operational Environmental Satellite - R series (GOES-R) is presented in this paper. A major challenge in the validation of the FSW product is the lack of ground truth flood maps and similar reference products from other satellite systems and other sources. To overcome this limitation, a two-level validation scheme for the FSW product is developed using the Moderate-resolution Imaging Spectroradiometer (MODIS) data as a proxy. In the first level, gauging station data collected by the U.S. Geological Survey (USGS) are employed as ground truth flood point information on local scales to verify the effectiveness of the proposed algorithm for flood detection. Gauging station data collected during 34 flood cases that occurred in 2010 and 2011 in the continental US were validated and assessed according to the rate of correct detection. Results showed that 79.71% of flooding stations were accurately detected from the MODIS 1 km images by the proposed FSW algorithm. In the second level of the validation, FSW detection results using the proposed algorithm were compared to the reference flood maps, which were generated by a supervised support vector machine (SVM) classification followed by human interpretation and editing. Flood detection accuracies for three major flood events occurred in Asia and Australia in 2010 were evaluated. Confusion matrices were employed as the accuracy measurement for the second level of the validation. Commission errors for the three flood cases were 6.75%, 13.45% and 21.45%, respectively. Omission errors of flood pixels varied between 9.58% and 19.61%. The validation results suggest that the employed FSW algorithm is capable of producing flood and standing water maps in an operational environment, and it meets the required accuracy and execution time of the product.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

The flood and standing water (FSW) product is an option 2 product in the Geostationary Operational Environmental Satellites - R series (GOES-R) program, which is the next generation of operational meteorological satellites to replace current GOES systems (NOAA, 2009). The GOES-R series satellites are in the development phase, and the first GOES-R series satellite is scheduled for launch in 2015 (NOAA, 2010a). The FSW product will employ reflectance data collected by the Advanced Baseline Imager (ABI) sensor to identify water pixels through a classification process in the presence of a cloud mask. In accordance with level one requirement of the GOES-R series, the required measurement accuracy of flood pixels will be 60% correct classification, and the refresh rate or coverage time will be 60 minutes (NOAA, 2009). The product will be only available during the daytime, which is achieved by restricting the solar zenith angle with a threshold

of 67 degrees. Cloud contaminated pixels will be omitted with the help of a cloud mask.

In the development of the GOES-R FSW algorithm and product, validation is a critical process. Validation provides information on the quantitative uncertainty required for the proper application of the product (Yu et al., 2011). Although flood identification studies based on optical remote sensors aboard polar-orbit or geostationary satellite platforms have been conducted for decades, e.g. Jain et al. (2006), Pantaleoni et al. (2007), Sakamoto et al. (2007), Sheng et al. (2001), the operational level of a flood and standing water product aboard a geostationary satellite has not been previously reported. Therefore, there are no heritage products available from other similar satellite systems that could be used as reference flood maps in the GOES-R FSW product validation.

Here, flooded areas are defined as normal land surfaces that are temporarily submerged in water. Compared to satellite-based measurements with a well-defined physical basis, such as land surface temperatures or soil moistures, the extent of a flood is usually difficult to be precisely surveyed using ground measurements. Hydrological gauging stations, which record the current and flooding stage of the monitored river or

* Corresponding author at: MS 6C3, 4400 University Drive, Fairfax, VA 22030, USA. Tel.: +1 703 993 5207; fax: +1 703 993 9230.

E-mail address: rzhang5@gmu.edu (R. Zhang).

stream, are the only available data source for indicating flood status at a given point of interest. By comparing the current stage and historical flood stage, gauging station data tell hydrologists whether or not the current location is experiencing a flood. Because flood status data obtained from gauging station observations represent precise measurements of river stages, they can be used as a trusted flood ground truth measurement data source in the validation of the satellite flood product. In this paper, gauging station flood status data obtained from the U.S. Geological Survey (USGS) were used. Brakenridge et al. (2012) first used gauging station data to validate satellite derived river floods. Here, a similar strategy for gauging station data was employed to validate the operational-level satellite derived flood and standing water product.

There are three limitations in validating a satellite flood product using the USGS gauging station data: 1) Gauging station data are point-based measurements that only indicates the flooding status for a specific point of interest, whereas the satellite-derived FSW product provides an area observation, which may cover a large area of lands submerged in water. This inconsistency in the characteristics of the two types of data hinders the proper validation for most flood affected areas. 2) Gauging stations are only located near rivers and streams, and only river flood events can be recorded by gauging stations. However, floods can occur at any location, including flash floods caused by heavy rains not associated with a river. On the contrary, these flood events can be observed by remote sensors, even though no gauging station data are available. 3) The gauging station data used in this study are obtained from the USGS, which only covers flood events that occurred inside the US territory. Floods that occurred outside of the US cannot be validated using the USGS gauging station data.

To overcome these limitations, flood pixels-based area validations for selected flood cases are also included in this study. This is the second level of the validation process, which is used to estimate the correct classification rate of the flood identification at area level. The characteristics of water during the daytime based on reflectance data are relatively easy to be distinguished from other land covers; hence the identification of the water pixels in images by experienced interpreters is usually accurate. However, direct manual extraction of the contours of the water bodies would be extremely time-consuming, and certain minor water bodies and isolated water pixels tend to be missed by human interpreters. Therefore, a two-step procedure that includes SVM classification and further human editing is implemented. A preliminary water-land classification map is generated by a supervised SVM classification process, and the classification map is further edited and revised in accordance with the visual interpretation. Finally reference flood maps are achieved and utilized in the quantitative validation of the FSW product.

In this paper, the validation process for the flood and standing water product is presented in detail; the C4.5 decision tree algorithm, which serves as the base classifier for the water-land separation, is also briefly introduced. Because the flood and standing water product algorithm will be run in an operational environment, the running time for processing the proxy data is also recorded. The remaining of the paper is organized as follows: In section two, the proposed flood and standing water detection algorithm is described. In section three, point validations using gauging station data are presented. The area validation that is based on the measurement of the flood detection accuracies is given in section four. Section five gives conclusion.

2. Data and FSW Algorithm

2.1. Data

Four data sets were used in this study: Moderate-resolution Imaging Spectroradiometer (MODIS) data obtained aboard Terra or Aqua platforms. MODIS data includes reflectance images; geolocation data and

cloud mask data, which are used as proxy for future ABI data. USGS gauging station stage data, the new MODIS 250 m land-water mask (short name: MOD44W), and global land cover data from Boston University.

MODIS data is used in the algorithm development because MODIS sensors are able to generate images with the same spatial resolution (1 km) of the future ABI sensors, and they possess similar reflectance channels to the ABI sensors. The wavelength ranges for all ABI bands are listed in the NOAA specification (2010b). In the algorithm development, the employed 0.64 μm band (channel 2) of ABI is simulated with the channel 1 of MODIS sensors, the 0.865 μm wavelength (channel 3) is simulated with the MODIS channel 2, and the 1.61 μm (channel 5) is simulated with the MODIS channel 6.

The gauging station records containing water stage and geolocation information were obtained from the US Geological Survey (USGS) Water Watch website (<http://waterwatch.usgs.gov/new/>). Flood status information was parsed and extracted from the webpages of this site, and the latitude and longitude of the flooding stations were also obtained for geolocation matching. The flood events that were used in the validation were confirmed using the NASA Earth Observatory. The flood cases used in the first level of the validation are listed in Table 1:

To obtain flood maps from the water-land classification maps, a reference non-flood water-land mask is required for comparing the classified results in a change detection process. In this study, the new MODIS 250 m land-water mask product (MOD44W) (Carroll et al. 2009), which was mainly created with Shuttle Radar Topography Mission (SRTM) Water Body dataset (SWBD), was used as the primary reference data. In addition, the global MODIS land cover data compiled by the MODIS land cover group at Boston University was employed as well (BU, 2009), as retrieved from ftp://crsftp.bu.edu/modis/MOD12Q1_data/. The land cover categories of the reference data follow the definitions of the International Geosphere-Biosphere Programme Land Cover Classification (IGBP) (Belward, 1996).

The MOD44W product is a significant improvement over other global surface water datasets (Carroll et al., 2009). In the MOD44W, the SWBD was used because of its fine spatial resolution and because of its consistent representation of the land surface. Because the SRTM data were collected over a short time period of only 11 days, it should provide a spatially coherent representation of surface water. Additionally, the cloud penetrating properties of the Radar offers superior performance over optical data alone, particularly in cloudy areas such as the humid tropics (Carroll et al., 2009). Although the MOD44W mask is a superior representation of land water over other datasets, the inconsistency in imaging quality between MODIS images and MOD44W mask may lead to extra detection errors due to the limitations of imaging capability of the MODIS sensors. To differentiate the errors from data sources and the FSW algorithm, the BU global land cover map, which was also generated from MODIS product representing the imaging capability of the MODIS sensors, was used as the secondary reference water mask in the validation. The details are described in Section 3. It is worth noting that the secondary reference data is only used in the gauging station validation, since the extra detection errors from the MOD44W are not statistical significance in the context of thousands of flood pixels in detection results in the second level of the validations.

2.2. Flood and standing water algorithm

2.2.1. Reflectance channels for water identification

To accurately identify water body pixels using remote sensing images, the spectral signatures of water and different ground cover types in different reflectance channels are investigated for the selection of the potential reflectance channels in the FSW algorithm. A reflectance spectra map for water and several land surface types is presented in Fig. 1.

Table 1

Summary of the used flood cases in the first level of the validation.

Date	State	Satellite	Date	State	Satellite	Date	State	Satellite
20100504	TN	T	20110429	TN	T	20110625	NE	T
20100509	KY	A	20110503	IN	T	20110627	MO	T
20100619	MO	T	20110504	IL	T	20110630	MO	T
20100624	SD	T	20110507	ND	T	20110704	MO	A
20110301	IN	T	20110515	MS	A	20110710	MO	T
20110312	IL	A	20110520	MO	T	20110718	NE	A
20110412	WI	T	20110601	MS	T	20110725	SD	A
20110414	ND	T	20110605	MT	A	20110801	SD	A
20110428	TN	T	20110606	NE	T	20110802	MO	T
20110429	AR	T	20110611	ND	T	20110814	MO	A
20110429	IN	T	20110615	NE	A			
20110429	KY	T	20110623	ND	T	Total:34		

Note:

TN: Tennessee; KY: Kentucky; MO: Missouri; SD: South Dakota; IN: Indiana; IL: Illinois; WI: Wisconsin; ND: North Dakota; AR: Arkansas; MS: Mississippi; MT: Montana; NE: Nebraska; T: Terra; A: Aqua.

Typical spectral reflectance curves of major ground-types during floods were previously studied in (Sheng et al., 1998). During floods, water body albedo increases significantly with a maximum reflectance peak moving towards the red band because silt and debris are concentrated in the water; in contrast, an increase in soil moisture causes a decrease in the soil albedo. Consequently, the reflectance characteristics of ground covers become complicated during floods, thereby preventing a simply distinction between water and land in individual AVHRR channels (Sheng et al., 1998). According to their study, water shows a lower albedo in the near-infrared band than other land types, whereas water pixels present higher reflectance characteristics in the red band than those of vegetation, soil and other ground cover types. Therefore, the ratio band of channel 2 (near-infrared) to channel 1 (red) of the AVHRR is employed to distinguish water and land.

Many other water body identification methods have also been developed. The thresholds of distinguishing between water and land pixels are based on the reflectance of the near-infrared channel (Sheng et al., 1998), the brightness temperature of channel 4 (10.3 μm) of the AVHRR (Barton and Bathols, 1989), and the difference between the near-infrared and red channels (Sheng et al., 1998). In the proposed GOES-R ABI flood and standing water detection algorithm, the red channel (channel 2), two near-infrared channel (channel 3) and (channel 5) of the ABI sensor, the difference of the two channels (channel 3-channel 2), the ratio of the two channels (channel 3/channel 2), and the simplified NDVI and revised Normalized Difference Water Index (NDWI) (Sun et al., 2012), which is inspired by the NDWI for open water (McFeeters, 1996), and the NDWI for leaf water content

(Gao, 1996), are employed as predictors to differentiate the water and land pixels. Additional details regarding the selection of predictors can be found in the algorithm theoretical basis document of the FSW product (Sun et al., 2011). In the development phase, the Channel 2, 3 and 5 of the ABI sensor are simulated using Channel 1, 2 and 6 of the MODIS sensors, respectively.

2.2.2. C4.5 Decision Tree Classifier

Because it is difficult to simultaneously tune multiple thresholds for the precise separation of the water and lands, the decision tree classifier, which is a series of well-established pattern recognition algorithms, is used to automatically determine the thresholds using a tree model construction in the training process. In the FSW algorithm, the classic C4.5 decision tree algorithm is employed as the basic classifier not only because it can achieve higher classification accuracy than simple classification rules (Elomaa, 1994), but it also typically runs faster during predictions than newer classifiers such as support vector machines, and artificial neural networks, which is critical for an operational algorithm.

C4.5 algorithm builds the decision tree using the concept of the information entropy and utilizes the normalized information gain (difference in entropy) as its criterion to choose attributes for splitting the data.

In pseudo code, the algorithm for building decision trees is as follows (Kotsiantis, 2007):

1. Check for base cases
2. For each attribute a
Find the normalized information gain from splitting on a
3. Let a_{best} be the attribute with the highest normalized information gain
4. Create a decision *node* that splits on a_{best}
5. Recur back to the sublists obtained by splitting on a_{best} , and add those nodes as children of *node*.

For a detailed explanation of the concept and theory of using the C4.5, see (Quinlan, 1992). A detailed description of the flood and standing water algorithm for the GOES-R ABI data can be found in the publication of the algorithm theoretical basis document (Sun et al., 2011).

For the training of the C4.5 decision tree, 13 MODIS flood cases and 17 MODIS non-flood cases acquired in 2009 and 2010 throughout the world were collected. To separate the water pixels from other ground types, the problem was modeled as a binary water or land classification. Water and land samples for the training process were manually selected from interactive views of the reflectance images. A total of 96,790 pixels were selected as input into the C4.5 algorithm

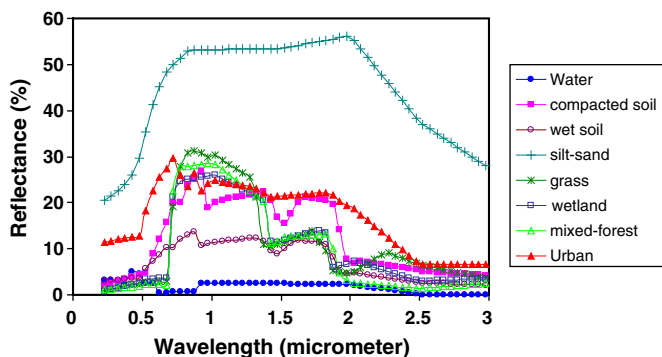


Fig. 1. Reflectance spectra of water and several other land surface types.

for the training process. The size of the tree (i.e., number of leaves) is 49. The training accuracy is 99.72% calculated by a 10-fold cross validation.

2.2.3. Post-classification change detection for flood identification

After separating the water and land pixels using the trained tree model in the classification process, the flooded areas are obtained according to the post-classification change detection using the non-flood reference data, in which the MOD44W data is employed. The classification is applied to the original MODIS L1b Swath data, and the classification result is compared to the non-flood reference data in a pixel-by-pixel manner in the same coordination system. After comparing the detected water-land classification results to the global land cover map in the post-classification change detection process, the pixels are divided into several categories: land-to-water (flood), water-to-land (drought), water-to-water (permanent water), and land-to-land (permanent land). Cloud affected pixels are omitted.

3. Gauging station point validation

34 flood cases that occurred in the US in 2010 and 2011 were validated using the USGS gauging station flooding data, and two flood cases were showed in detail in this section. The flood status (i.e., flood or non-flood) of a gauging station was obtained by comparing the current stage to the flood threshold stage. It is worth noting that the gauging level mainly indicates the flood status in the downstream areas. The geolocations of collected stations were reprojected into a uniform coordination system for match-ups with MODIS images. In this study, latitudes and longitudes in North American Datum of 1927 (NAD27) or NAD83 were reprojected onto the Universal Transverse Mercator (UTM) projection zone 14 N. In addition, the MODIS reprojection tool-swath was utilized to reproject MODIS 1 km L1b Swath data, the MOD44W non-flood reference data and FSW detection results onto the same coordination system (USGS, 2010).

Generally, pixel distortions that are introduced in reprojection process are unavoidable, especially in edge areas. In addition to systematic geometric errors that occur in the MODIS L1b data, there can be considerable error in the geolocation of a pixel. Moreover, the spatial resolution of a pixel of the reprojected images (approximately 0.01 degree) and the positioning accuracy for a gauging station, which is obtained by highly accurate ground or GPS surveys (approximately 0.1 second), are inconsistent. Due to these factors, it is usually not possible to match a single gauging station to an exact pixel in 1 km remote sensing images. Therefore, a buffer zone containing multiple pixels is used in the images to match the geolocation of the gauging station. In this study, a 3×3 pixel rectangular box is applied in the images. Because flood detection results, the reference data and the station coordinates are reprojected into the same coordination system using identical projection parameters, the inconsistency of accuracy in geolocations of image data and station coordinates is minimized, that enable us to obtain accurate pixel to station match-ups.

If at least one pixel is identified as flood in the 3×3 box, and the corresponding gauging station data confirm that the station is flooding, then the flood location in the image for the station is recognized correctly.

If there are flood pixels in the box, but the station indicates no flooding, then the detected flood location in the station can be recognized as a commission error. If no flood pixel is detected, while the station indicates that it is flooding, the detected result for the station is recognized as an omission error. Cloud affected pixels are omitted.

Although a gauging station can be built in any large river or small stream, sections of small streams cannot be observed using satellite sensors due to the spatial resolution capability of the sensor. Here, only the rivers and streams possessing greater than 1 km width can be detected by the sensor and the FSW algorithm. Stations used to monitor unobservable small streams cannot be detected using MODIS 1 km images. Because the primary reference data of the MOD44W is generated in much higher resolution from the SRTM radar sensor, the aggregated MOD44W reference data will present part of water pixels representing minor streams that cannot be directly observed in MODIS 1 km images. That would introduce false omission errors in flood detection maps, which are actually not observable due to the imaging capability of the sensors. To exclude these unobservable stations, the secondary reference non-flood data from the BU global land cover map is employed to locate unobservable water pixels. Because the Boston University global land cover map is also generated using the MODIS sensors, water pixels marked by this data indicate the rivers or other water bodies that are observable and detectable based on the sensor and algorithm. In the flood and standing water detection results, a pixel that does not indicate drought or water suggests that no water pixel is recorded in the corresponding location in the BU land cover map, and if the pixel is recognized as an omission, it will be re-labeled as not observable in this step. Finally, correct detections, commission errors and omission errors for stations derived from the stations-to-areas match-up validation represent the effective gauging stations in the validation; other stations recording no floods and unobservable stations are ignored. All effective stations will be used in the final accuracy statistics.

3.1. Case 1: Floods in Illinois caused by levee breach

After the U.S. Army Corps of Engineers breached a levee to save the city of Cairo, Illinois, on May 2, 2011, agricultural fields remained flooded two days later. The levee breach flooded roughly 130,000 acres of nearby farmland in the Birds Point-New Madrid Floodway (NASA, 2011b; NOAA, 2011a).

MODIS Terra data acquired on May 4, 2011 were used as the after flood image to identify floods, and Terra data acquired on April 29, 2010 were also collected as the baseline image for visual comparisons. All gauging station data in the Illinois state were collected and used in the station validation. With the exception of cloud covered stations, five representative gauging stations recording the flood event are listed in Table 2:

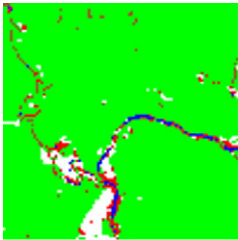
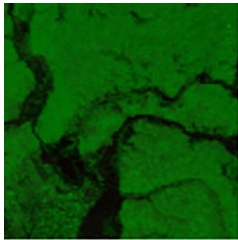
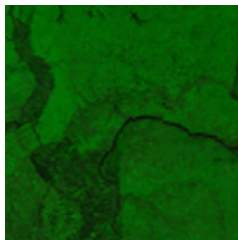
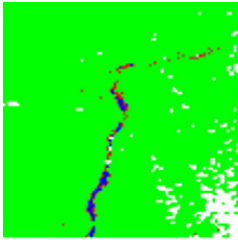
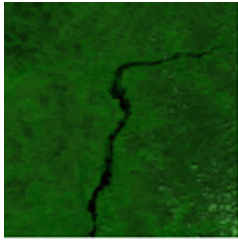
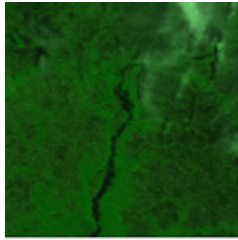
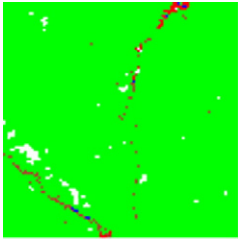
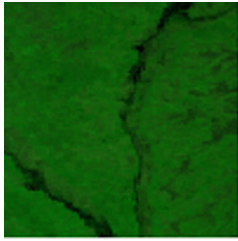
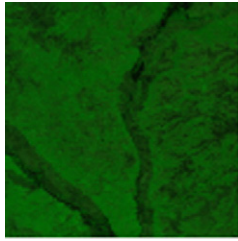
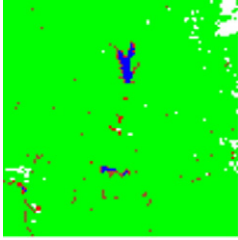
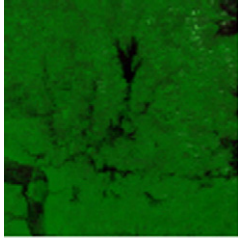
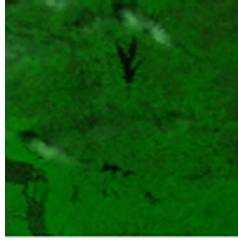
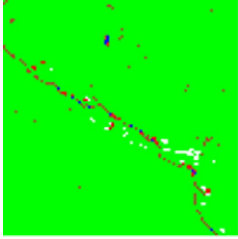
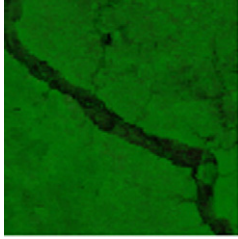
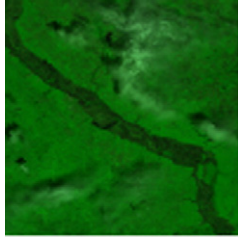
The subsets of the detected flood maps, the after flood and before flood comparison images and the analysis results obtained by the above mentioned strategy are listed in Table 3. Subsets of 100×100 pixels for each station are illustrated for comparison, and the tested gauging stations are located approximately in the center of each image. In the flood maps, the red pixels are flooded areas, the greens

Table 2
Gauging station information for the Illinois flood case.

No.	Station Code	Name	Stage (ft)	Flood stage (ft)	date	Latitude	Longitude
1	03612500	OHIO RIVER AT DAM 53 NEAR GRAND CHAIN, IL	58.19	42	2011-05-04	37°12'11"	−89°2'30"
2	05558300	ILLINOIS RIVER AT HENRY, IL	25.28	23	2011-05-04	41°6'26"	−89°21'22"
3	05586100	ILLINOIS RIVER AT VALLEY CITY, IL	20.05	14	2011-05-04	39°42'12"	−90°38'43"
4	05597000	BIG MUDDY RIVER AT PLUMFIELD, IL	34.68	20	2011-05-04	37°54'5"	−89°0'50"
5	07020500	Mississippi River at Chester, IL	37.78	27	2011-05-04	37°54'13.5"	−89°50'8"

Table 3

Flood detection results and comparisons for the Illinois flood case.

No	Flood map	After flood image	Pre flood image	Result
1				Omit
2				Correct
3				Not detectable
4				Correct
5				Correct

are land, and the permanent water bodies are marked in blue. The white pixels are clouds, which are provided by the MOD35 cloud mask product. It is worth noting that certain water pixels are incorrectly marked as clouds; this error is introduced by the cloud mask. In this study, no further processes are performed on these incorrect pixels.

3.2. Case 2: Flooding in the Missouri Basin

The Missouri River remained flooded along the borders between South Dakota, Iowa, and Nebraska in late July 2011. The MODIS on NASA's Aqua satellite captured the flood event on July 25, 2011. For

comparison, Aqua data in the same area a year earlier, on July 17, 2010 was also acquired (NASA, 2011a). All gauging station data in South Dakota were collected and used in the station validation. The same procedure is applied to this flood event; the gauging station information and the flood detection subset images are summarized in Tables 4 and 5, respectively.

The proposed validation strategy was conducted for all collected flood cases, which are summarized in Section 2, and a total of 69 effective gauging stations were obtained. For all effective stations, 55 flooding stations were correctly detected in the images using the FSW algorithm, 2 stations were identified as commission errors, and

Table 4

Gauging station information for the Missouri Basin flood.

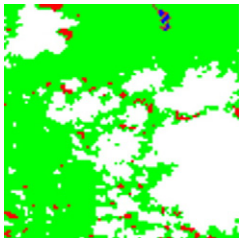
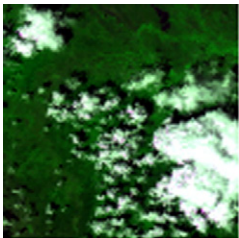
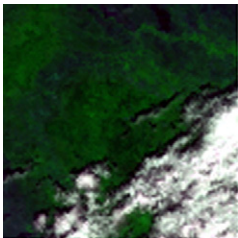
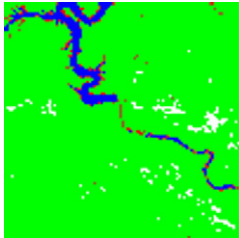
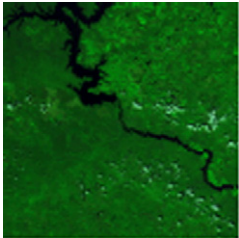
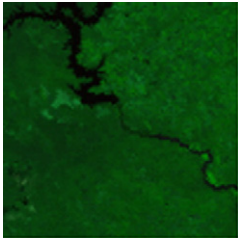
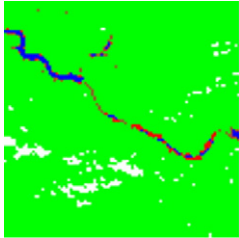

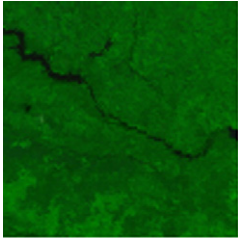
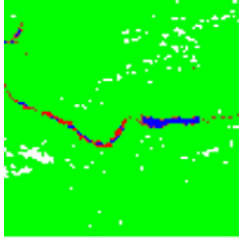

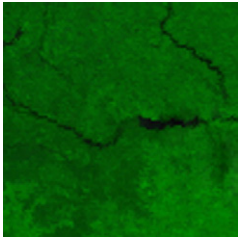
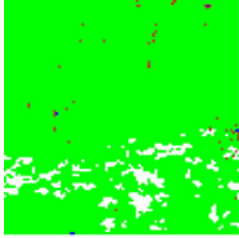
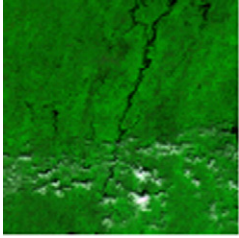
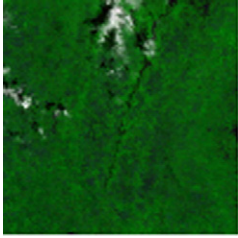
No.	Station code	Name	Stage (ft)	Flood stage (ft)	Date	Latitude	Longitude
1	06430800	ANNIE CREEK NEAR LEAD, SD	4.19	10	2011-07-25	44°19'39.05"	–103°53'38.58"
2	06440000	MISSOURI R AT PIERRE,SD	18.29	13	2011-07-25	44°22'23"	–100°22'3"
3	06453020	MISSOURI RIVER BELOW GREENWOOD, SD	38.72	30	2011-07-25	42°54'19"	–98°20'58"
4	06466700	LEWIS AND CLARK LAKE AT SPRINGFIELD,SD	14.16	10	2011-07-25	42°51'21"	–97°53'6"
5	06472000	JAMES R NEAR STRATFORD,SD	21.1	14	2011-07-25	45°14'30"	–98°23'28"

12 stations were omitted. When divided by the total number of effective gauging stations, the correct detection rate is 79.71%, the commission error rate is 2.90%, and the omission error rate is 17.39%.

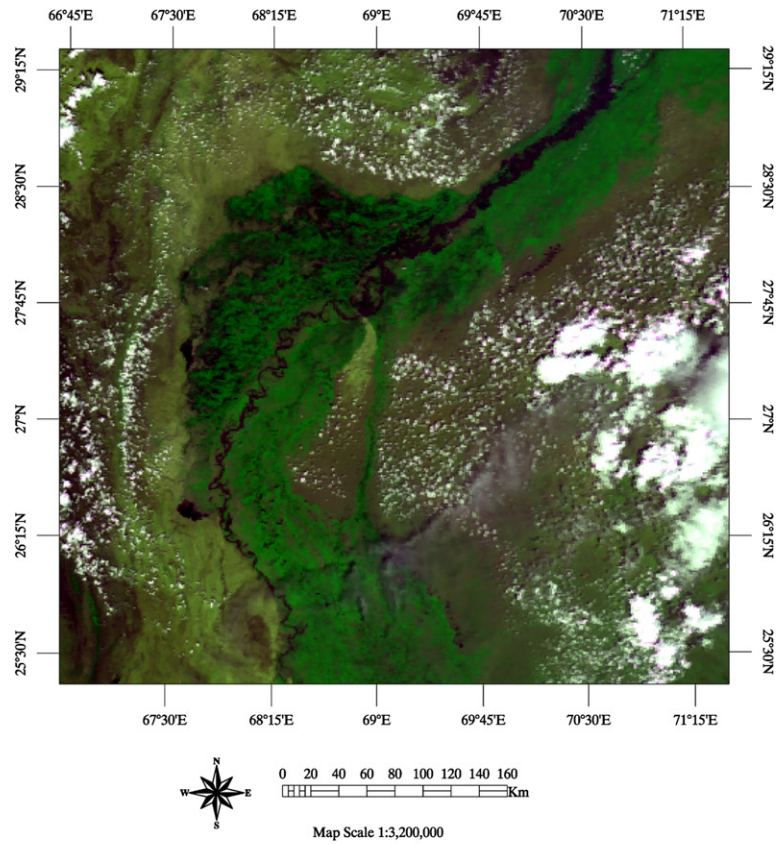
Based on the quantitative evaluation using the correct detection rates, the obtained 79.71% of flood event detection accuracy exceeds the classification accuracy requirement of 60% for the GOES-R product.

Table 5

Flood detection results and comparisons for the Missouri Basin flood.

No	Flood map	After flood image	Pre flood image	result
1				Commit
2				Omit
3				Correct
4				Correct
5				Not detectable

a) Before flood subset image view acquired at 09:00 on August 28, 2009



b) After flood subset image view acquired at 08:30 on August 28, 2010

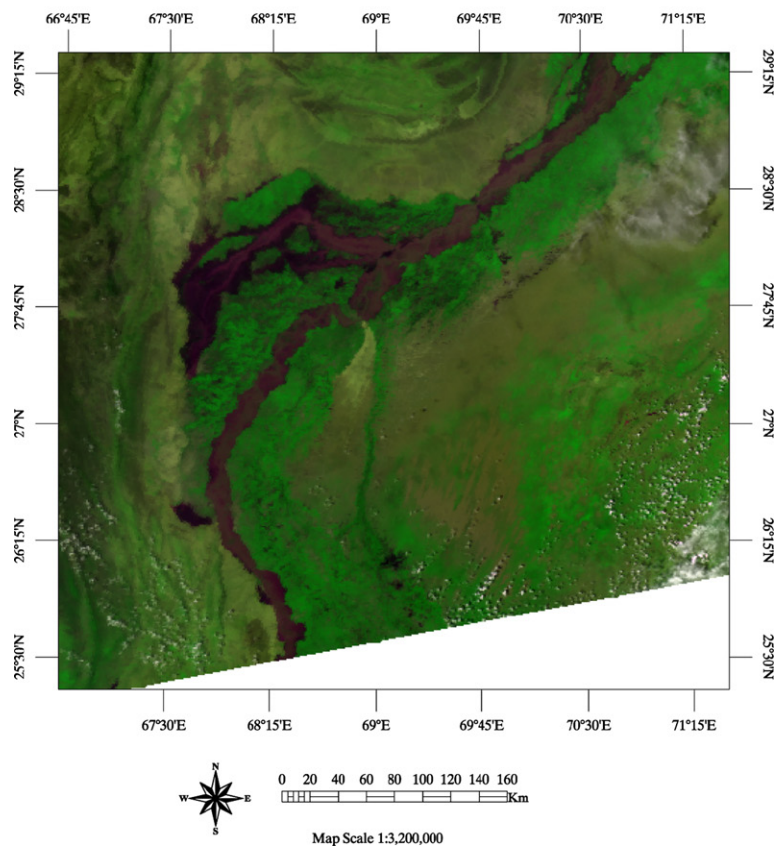


Fig. 2. MODIS Aqua reflectance composite image view for the Balochistan flood. Channel 1 (R), channel 2 (G), channel 3 (B) is shown in composite.

The algorithm presents a very low commission error rate, which indicates that the proposed algorithm seldom falsely recognize non-flood pixels as floods in local scales. When comparing of after and before flooding images, occasional commission errors are caused by confusions between cloud shadows and water pixels, such as those observed for station 1 in case 2. Because the cloud shadow mask is not a basic input in the operational phase, and a reliable cloud shadow product is not available for MODIS products, a cloud shadow is not applied in the validations. The FSW software package has the ability to apply an additional cloud shadow mask to improve the detection accuracy, although this is an optional input for the software. The results show a relatively larger omission error than the commission error for all tested cases, which indicate that the proposed algorithm may tend to miss some flood areas, for example station 1 in case 1 and station 2 in case 2. The validation also gives two “not detectable” stations using 1 km MODIS data for demonstration, such as station 3 in case 1 and station 5 in case 2.

4. Validation of the interpreted images

This section presents the image-level validations performed using the confusion matrices and the detection accuracies of three selected major flood cases that occurred in 2010. The detected flood maps are compared with the reference flood maps, and the flood mapping accuracies based on correctly detected flooding pixels are measured. Because there are no similar flood products available from other satellite systems, the reference flood maps used in the comparisons are also generated in this study by classification and visual interpretation. To ensure the accuracy of the locations of the true flood pixels, the reference flood maps

Table 6

Confusion matrix for accuracy for the Balochistan flood.

Class	Flood	Water	Land	Total
Flood	9458	105	580	10143
Water	106	267	14	387
Land	1265	14	1674001	168679
Total	10829	386	167994	179209

Overall accuracy = $(177125/179209)$ 98.8371%.

Kappa coefficient = 0.8982.

Flood commission error = $(685/10143)$ 6.75%.

Flood omission error = $(1371/10829)$ 12.66%.

are generated in two steps. First, a support vector machines classifier is employed on 250 m MODIS reflectance images, in which regions of interest of samples are deliberately collected from each flood case for training, to generate an initial water-land classification map. Then flood maps are generated using the post-classification change detection process comparing with 250 m resolution MOD44W product, which has been described in Section 2. Second, the SVM-derived flood maps are manually edited according to the visual interpretation in an interactive environment. Some falsely classified pixels are assigned to new class labels. In this step, high-resolution Landsat Thematic Mapper images are used as the references to verify the locations of the water bodies. Compared to the direct extraction of the contours of water bodies, the proposed two step method reduces the human's workload, and small water bodies that are difficult for interpreters to identify can be discovered by the pixel classifier. Finally, the generated 250 m reference flood maps are aggregated into 1 km resolution for accuracy evaluations.

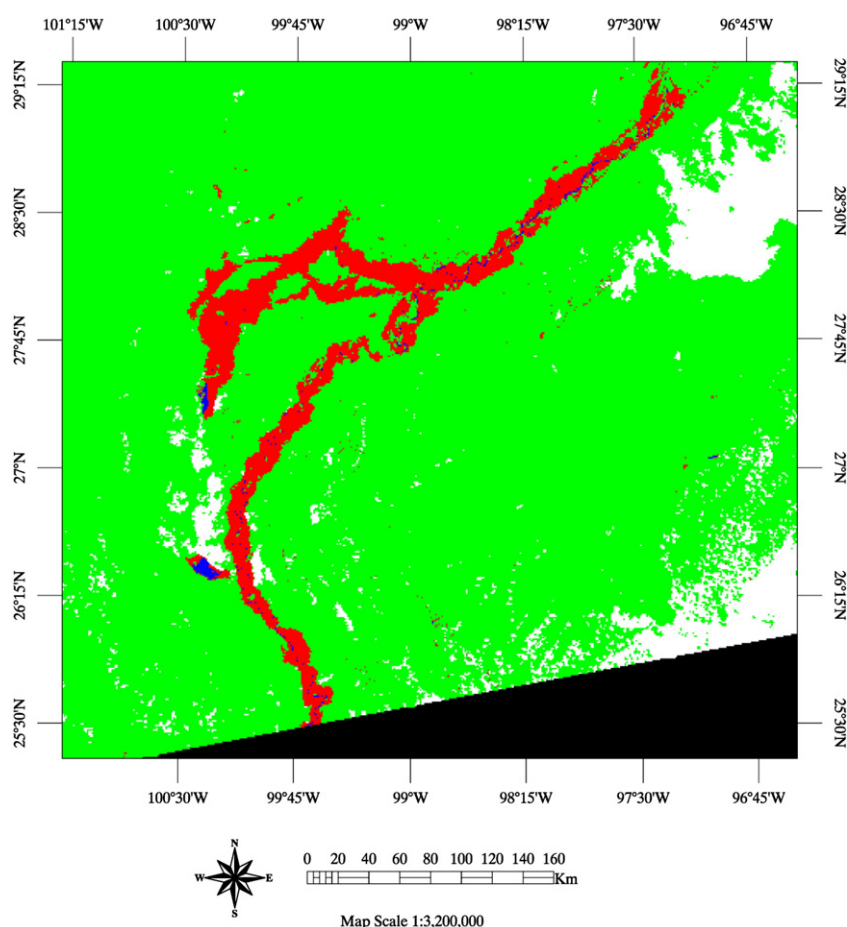
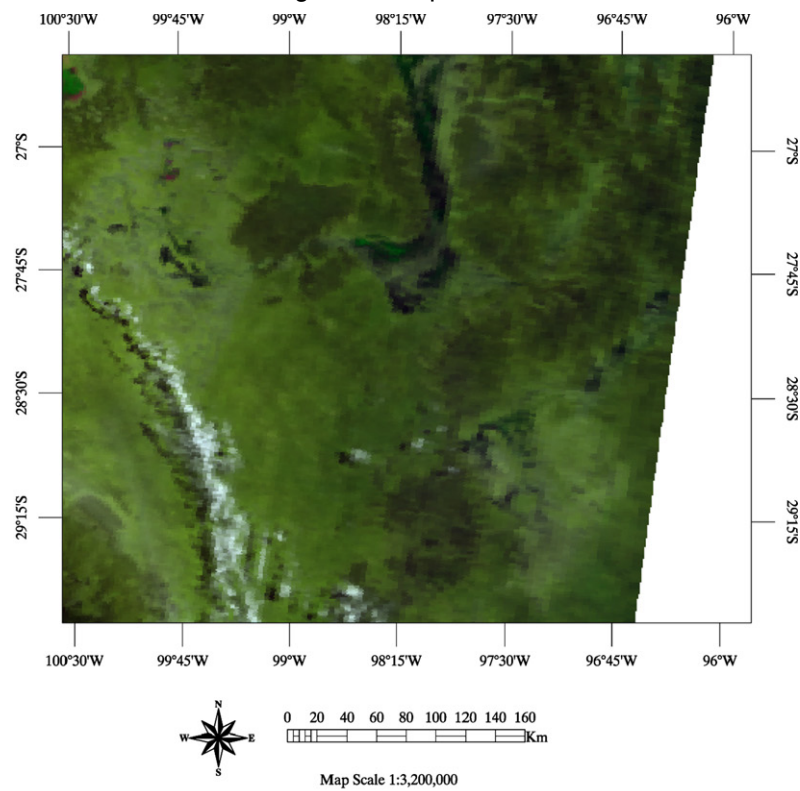


Fig. 3. Color-coded flood detection map for the Balochistan flood. Flood (red), land (green), permanent water (blue), and cloud (white) are displayed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The SVM classifier has been used extensively in numerous fields of research, including remote sensing (Huang et al., 2008, 2011; Zhang and Ma, 2008; Zhu and Blumberg, 2002). It can provide reliable classification

results for remote sensing applications, which has been confirmed by various studies. Detailed descriptions and explanations of the concept of SVM can be found in Burges (1998); Cristianini and Shawe-Taylor

a) Before flood subset image view acquired at 01:25 on March 26, 2009



b) After flood subset image view acquired at 00:55 on March 26, 2010

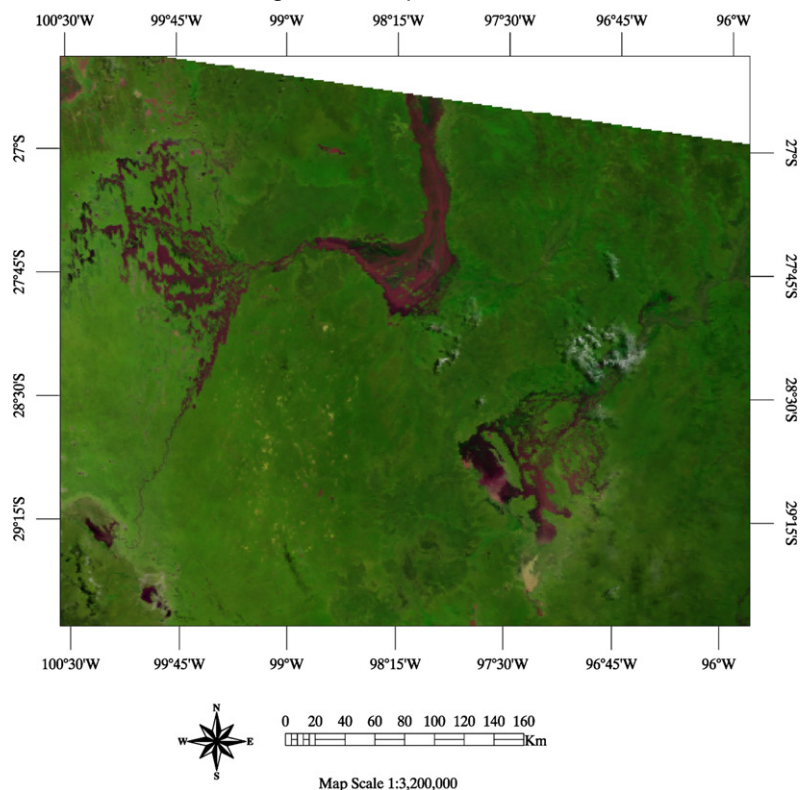


Fig. 4. MODIS Terra reflectance composite image view for the channel country flood case. Channel 1 (R), channel 2 (G), channel 3 (B) is shown in composite.

(2000); Vapnik (2000). Compared to traditional classifiers, the SVM does not assume distributions of the input data, as does maximum likelihood, and it does not require a complicated process for parameters tuning, such as in neural networks. In addition, the decision tree classifier is not selected in this step to avoid possible correlations between the detected results using the decision tree based algorithm and the generated true flood maps.

Because only the red channel and one near-infrared channel are presented in 250 m MODIS images, five predictors are involved in the classification: channel 1, channel 2, channel 2 - channel 1, channel 2/channel 1, and the NDVI. To ensure mapping accuracies, MODIS 250 m reflectance images and MOD44W data were projected into the UTM coordination system. For the tested case 1 and case 3, UTM zone 42 N was applied, and UTM zone 54 S was used in case 2.

The confusion matrix (Congalton and Mead, 1983; Kohavi and Provost, 1998) was employed as the measure of classification accuracy, and the flood, permanent water, and land classes were included in the confusion matrices. The commission and omission errors of the detected flood pixels were also presented. To reduce the statistical bias, only cases that possess more than or approximately 10000 detected flood pixels were selected in the validations.

4.1. Case 1: Flooding in Balochistan

In the summer of 2010, a serious flood from the Indus River, which flows through the provinces of Punjab and Sindh, hit the Balochistan Province of Pakistan (NASA, 2010a). 1 km MODIS Aqua data acquired on August 28, 2010, were used to detect this flood. The non-flood image acquired on August 28, 2009, was also collected for comparison. Before and after flood reflectance images are presented in Fig. 2 (a) and (b), respectively, and the flood detection map is shown in Fig. 3, in which the

red pixels denote floods, green areas are land, permanent water bodies are marked in blue, and clouds are white. The confusion matrix is presented in Table 6.

4.2. Case 2: Flooding in Channel Country

Floodwaters traveled southward in Australia's Channel Country in late March, 2010. Earlier in the month, flooding forced authorities to declare most of Queensland a natural disaster area and caused hundreds of millions of dollars' worth of damage. By the end of the month, floodwaters filled channels in the south while receding from some areas to the north (NASA, 2010b).

1 km MODIS Terra data acquired on March 26, 2010, were used to detect this flood. The non-flood image acquired on March 26, 2009, was also collected for comparison. Before and after flood reflectance images are presented in Fig. 4 (a) and (b), respectively, and the flood detection map is shown in Fig. 5. The confusion matrix is presented in Table 7.

4.3. Case 3: High waters along the Pakistan-India Border

Just inland from the Arabian Sea, wetlands line the border between Pakistan and India. Water levels rise and fall with the seasons, but after weeks of devastating monsoon rains that displaced millions of Pakistanis, water levels in the typically shallow lakes stayed high in late summer 2010 (NASA, 2010c).

1 km MODIS Terra data acquired on September 23, 2010, were used to detect this flood. The non-flood image acquired on September 23, 2009, was also collected for the comparison. Before and after flood reflectance images are presented in Fig. 6 (a) and (b), respectively, and

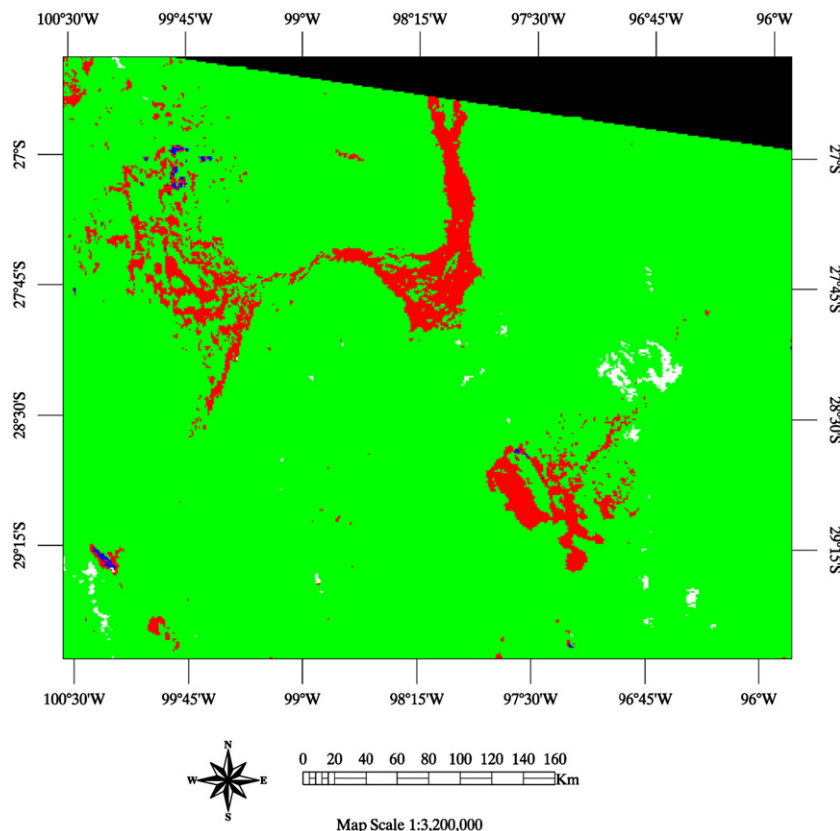


Fig. 5. Color-coded flood detection map for the channel country flood case. Flood (red), land (green), permanent water (blue), and cloud (white) are displayed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7

Confusion matrix for accuracy for the Channel Country flood.

Class	Flood	Water	Land	Total
Flood	8371	38	1263	9672
Water	26	84	8	118
Land	2016	12	164179	166207
Total	10413	134	165450	175997

Overall accuracy = $(172634/175997)$ 98.0892%.

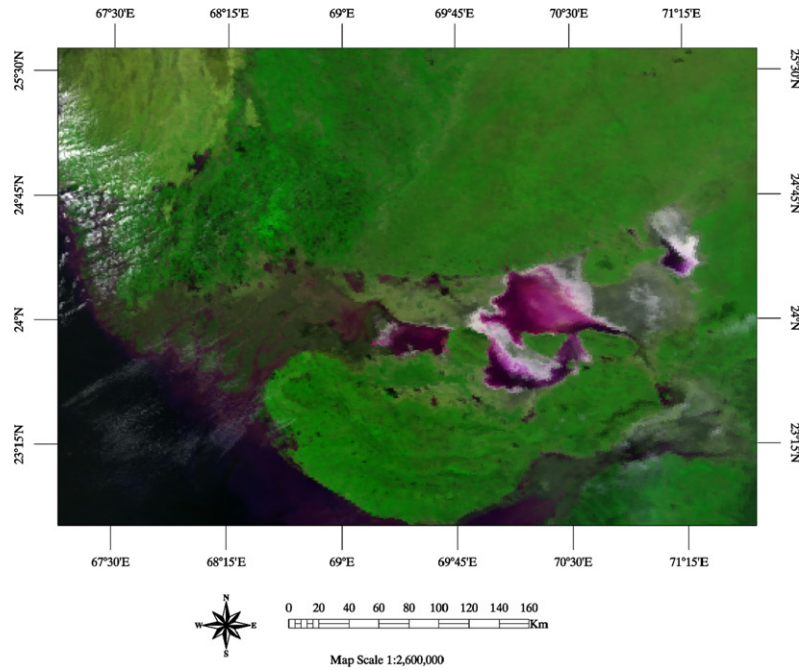
Kappa coefficient = 0.8246.

Flood commission error = $(1301/9672)$ 13.45%.Flood omission error = $(2042/10413)$ 19.61%.

the flood detection map is shown in Fig. 7. The confusion matrix is presented in Table 8.

Based on the confusion matrices for the three validated flood cases, all results show high total accuracies. Overall accuracies of 97.47% to 98.84% of indicate that almost all detected pixels are consistent with those in the reference flood maps. Kappa coefficient values varied between 0.8246 and 0.9353, which means that the algorithm predicts high classification accuracy for all three classes with no obvious classification bias. Because the kappa coefficient takes into account the agreement that occurs by chance, it is generally perceived as a more stable measure than a simple percent agreement calculation. The commission

a) Before flood subset image view acquired at 06:25 on September 23, 2009



b) After flood subset image view acquired at 05:55 on September 23, 2010

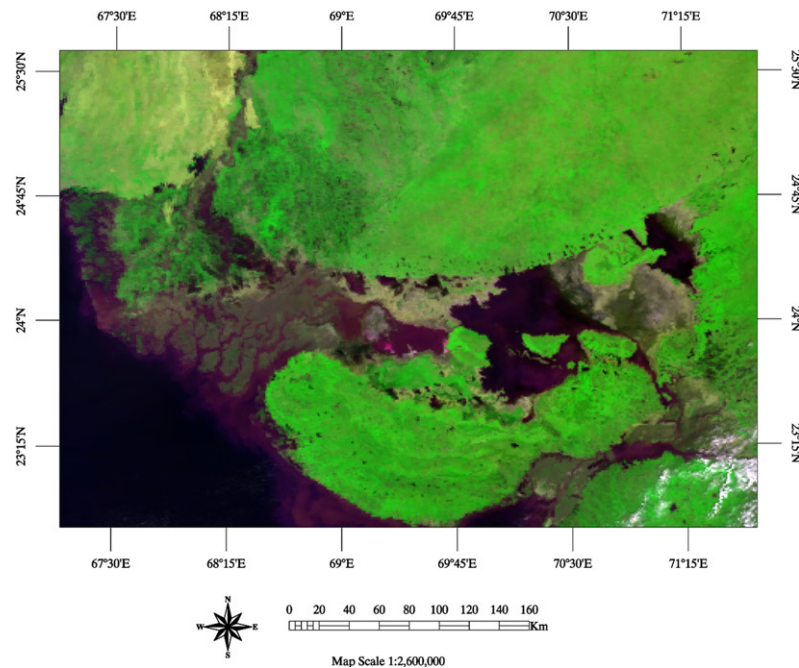


Fig. 6. MODIS Terra reflectance composite image view for the Pakistan-India Border flood case. Channel 1 (R), channel 2 (G), channel 3 (B) is shown in composite.

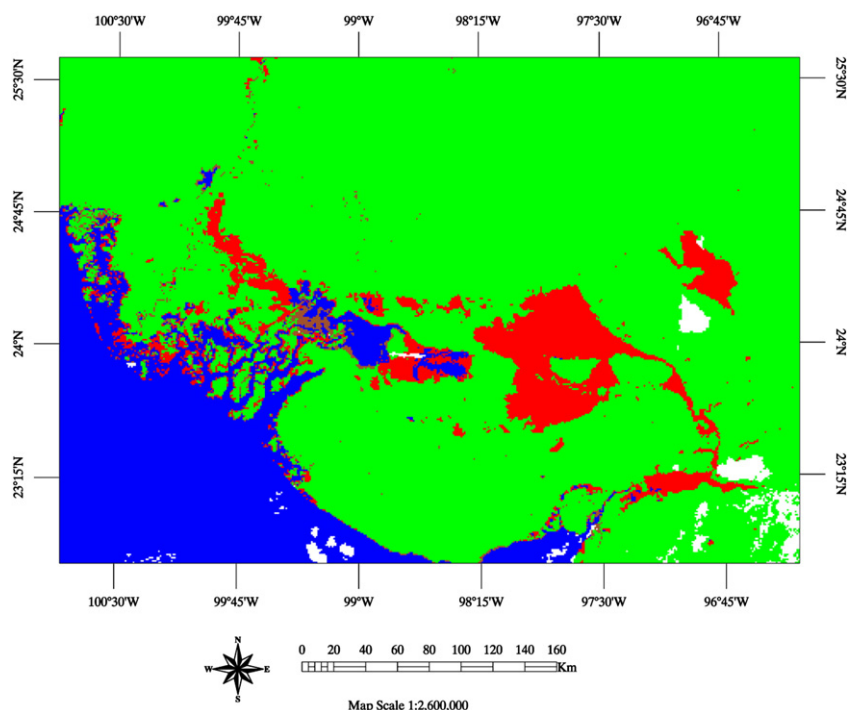


Fig. 7. Color-coded flood detection map for the Pakistan-India Border flood case. Flood (red), land (green), permanent water (blue), and cloud (white) are displayed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

errors in the validations vary between 6.75% and 21.45%, which indicates that flood pixels are misclassified into other classes. The omission errors for three flood cases are 12.66%, 19.61% and 9.58%, respectively. Both commission and omission errors for the test cases are considerably lower than the required 40% incorrect classification accuracy for flood identifications. The three area validation results of flood mapping accuracies suggest that the proposed FSW algorithm also meets requirements of flood detection accuracy.

Several factors could explain the obtained commission and omission errors. 1) Due to the limitation of the sensor's capabilities, a portion of water bodies covering small areas are difficult to be distinguished from 1 km remote sensing images. 2) During flood events, a portion of water pixels show relatively high reflectance values owing to different water constituents, such as high debris concentration. These flood water pixels tend to be mixed with lands. 3) The change of soil moisture in land surfaces close to water bodies during flood events makes it difficult to separate land pixels with water because of reflectance changes. 4) Due to the tradeoff between the final classification accuracy and the complexity of the decision tree model, the applied tree structure is pruned and the decision rules are simplified, which may introduce misclassifications for some flood pixels in predictions. There are two main reasons to perform tree pruning. One is to avoid possible model overfitting, which is typically caused by a very complicated training model, and then to achieve a better generalization ability. Second is to reduce

the classification time, which is proportional to the complexity of the applied model.

In these experiments, the running time of the algorithm was also recorded. Because the prediction process of a decision tree classification is relatively simple, the proposed procedure is fast. In the test cases, the program is able to produce flood detection maps within 15 seconds when MODIS proxy data is used as the input, which is much faster than the required 60 minutes in the operational environment. Even though the volume of a fulldisk image of future ABI data is larger than a MODIS 1 km granule image, the program is predicted to generate the product in 10 minutes. The running time measurements demonstrate that the proposed algorithm meets the computational requirement of the flood and standing water product.

The proposed FSW detection algorithm has also been utilized in real world applications. A flash flood caused by a tsunami, which followed the massive earthquake that struck northeastern Japan on March 11, 2011, was analyzed by the FSW algorithm as a rapid response to that disaster (NOAA, 2011b). The preliminary detection results for this flood event showed that most tsunami affected areas were accurately identified through a visual inspection. This case verified the effectiveness and usability of the proposed FSW algorithm in a real flood application.

5. Conclusion

In this paper, the validation process for the GOES-R flood and standing water product was described. The proposed algorithm using the C4.5 decision tree classification algorithm was also briefly described. The algorithm was validated using the MODIS 1 km reflectance data as the proxy in the development phase. Two types of validations, a station point validation and an image area validation, were also included in this study.

In the first level of validation, the sensitivity of the selected channels and the applicability of the proposed algorithm were validated using gauging station data. 79.71% of the flood events detection rate calculated on 34 flood cases was achieved. The validation results indicate that the selected predictors and classification algorithm are capable of detecting floods based on 1 km MODIS data. In the second level of the validation

Table 8
Confusion matrix for accuracy for the Pakistan-India Border flood.

Class	Flood	Water	Land	Total
Flood	8782	332	2066	11180
Water	231	23699	229	24159
Land	699	63	107091	107853
Total	9712	24094	109386	143192

Overall accuracy = $(139572/143192)$ 97.4719%.

Kappa coefficient = 0.9353.

Flood commission error = $(2398/11180)$ 21.45%.

Flood omission error = $(930/9712)$ 9.58%.

process, accuracy evaluations against classified and visual interpreted reference flood maps were performed. The commission and omission errors of the detected flood pixels were less than 25% for all tested cases, which are significantly less than the required 40% incorrect classification accuracy.

Not only the flood event detection accuracy that presented in the first level of the validation, but the flood mapping accuracy given in the second level of the validation indicates that the proposed FSW algorithm for the ABI sensors meets the mission's requirement of the 60% correct classification rate. For all test cases in this study, the proposed algorithm shows the effectiveness and robustness in tasks of flood detection and mapping visually and quantitatively. The FSW algorithm is built on a flexible pattern recognition classification framework, and detailed decision rules are obtained by training large samples of water and land collected around the globe, which represent variations of observed reflectance of ground types. Additionally, the decision model, which is stored in a tree structure, will be updated along with routine validations with new data. All these efforts are aimed at the usability of the algorithm for real flood conditions, and we have reason to believe the proposed algorithm is able to provide sufficient detection ability for a variety of real flood cases. The application in the Japan tsunami flood case, also confirmed the real world usability of the proposed FSW algorithm.

In the algorithm validation phase, the proposed two level validation strategies will be performed periodically. Moreover, comparisons against independent flood datasets, for example, the Dartmouth flood observatory (Brakenridge, 2010), will be involved in further developments of the validation system.

Acknowledgement

The authors would like to thank the reviewers for their valuable comments. This study was supported through the NOAA GOES-R Algorithm Working Group. The manuscript contents are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U.S. Government.

References

- Barton, I. J., & Bathols, J. M. (1989). Monitoring Floods with AVHRR. *Remote Sensing of Environment*, 30, 89–94.
- Belward, A. S. (Ed.). (1996). *The IGBP-DIS Global 1 km Land Cover Data Set "DISCOVER": Proposal and implementation plans*. Report WP No. 13, IGBP-DIS, Stockholm, Sweden.
- Brakenridge, G. R. (2010). The flood observatory: Space-based Measurement and Modeling of Surface Water. Available at: <http://floodobservatory.colorado.edu> (Accessed March 20, 2011).
- Brakenridge, G. R., Cohen, S., Kettner, A. J., De Groeve, T., & Nghiem, S. V. (2012). Satellite river discharge measurements. Available at: <http://floodobservatory.colorado.edu/SiteDisplays/IndexMap.htm> (Accessed Jan 16, 2012).
- BU (.). MODIS Land Cover Group. Available at: <http://www-modis.bu.edu/landcover/> (Accessed: Jan. 14, 2011).
- Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2, 121–167.
- Carroll, M. L., Townshend, J. R., DiMiceli, C. M., Noojipady, P., & Sohlberg, R. A. (2009). A new global raster water mask at 250 m resolution. *International Journal of Digital Earth*, 2, 291–308.
- Congalton, R. G., & Mead, R. A. (1983). A quantitative method to test for consistency and correctness in photo-interpretation. *Photogrammetric Engineering and Remote Sensing*, 49, 69–74.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines : and other kernel-based learning methods*. Cambridge;New York: Cambridge University Press.
- Elomaa, T. (1994). In Defense of C4.5: Notes on Learning One-Level Decision Trees. *Proc. of the 11th Int. Conf. on Machine Learning* (pp. 62–69) : Morgan Kaufmann.
- Gao, B. C. (1996). NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58, 257–266.
- Huang, L., Li, Z., Tian, B. -S., Chen, Q., Liu, J. -L., & Zhang, R. (2011). Classification and snow line detection for glacial areas using the polarimetric SAR image. *Remote Sensing of Environment*, 115, 1721–1732.
- Huang, C., Song, K., Kim, S., Townshend, J. R. G., Davis, P., Masek, J. G., et al. (2008). Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment*, 112, 970–985.
- Jain, S. K., Saraf, A. K., Goswami, A., & Ahmad, T. (2006). Flood inundation mapping using NOAA AVHRR data. *Water Resources Management*, 20, 949–959.
- Kohavi, R., & Provost, F. (1998). Glossary of Terms. *Machine Learning*, 30, 271–274.
- Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. In I. Maglogiannis, K. Karpouzis, M. Wallace, & J. Soldatos (Eds.), *Emerging Artificial Intelligence Applications in Computer Engineering - Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies* (pp. 3–24) : IOS Press.
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17, 1425–1432.
- NASA (2010). Flooding in Balochistan. Available at: <http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=45621> (Accessed: March 20, 2011).
- NASA (2010). Flooding in Channel Country. Available at: <http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=43288> (Accessed: March 20, 2011).
- NASA (2010). High Waters along the Pakistan-India Border. Available at: <http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=46030> (Accessed: March 20, 2011).
- NASA (2011). Flooding in the Missouri Basin. Available at: <http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=51472> (Accessed: Nov. 1, 2011).
- NASA (2011). Spring Floods in the U.S. Midwest and Canada. Available at: <http://earthobservatory.nasa.gov/NaturalHazards/view.php?id=50457> (Accessed: Nov. 1, 2011).
- NOAA (2009). GOES-R Series Level I Requirements (LIRD). Available at: http://www.goes-r.gov/overview/P417-R-LIRD-0137_v2_0_9-01-09.pdf (Accessed: December 20, 2010).
- NOAA (2010). GOES-R Home. Available at: <http://www.goes-r.gov> (Accessed: October 26, 2010).
- NOAA (2010). ABI-tech-summary. Available at: <http://www.goes-r.gov/spacesegment/ABI-tech-summary.html> (Accessed: October 26, 2010).
- NOAA (2011). The GOES-R Algorithm Working Group Analyzes Satellite Views of the Tsunami and Flood in the Aftermath of the Japanese Earthquake. Available at: http://www.star.nesdis.noaa.gov/star/news2011_201103GOESRtsunami.php (Accessed: March 18, 2011).
- NOAA (2011). Flooding in Wake of Levee Breach Seen through Next Generation Algorithm. Available at: <http://www.nvnl.noaa.gov/MediaDetail.php?MediaID=725&MediaTypeID=1> (Accessed: May 20, 2011).
- Pantaleoni, E., Engel, B. A., & Johannsen, C. J. (2007). Identifying agricultural flood damage using Landsat imagery. *Precision Agriculture*, 8, 27–36.
- Quinlan, J. R. (1992). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Series in Machine Learning: Morgan Kaufmann.
- Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., & Yokozawa, M. (2007). Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery. *Remote Sensing of Environment*, 109, 295–313.
- Sheng, Y., Gong, P., & Xiao, Q. (2001). Quantitative dynamic flood monitoring with NOAA AVHRR. *International Journal of Remote Sensing*, 22, 1709–1724.
- Sheng, Y. W., Su, Y. F., & Xiao, Q. G. (1998). Challenging the cloud-contamination problem in flood monitoring with NOAA/AVHRR imagery. *Photogrammetric Engineering and Remote Sensing*, 64, 191–198.
- Sun, D., Yu, Y., Zhang, R., Li, S., & Goldberg, M. (2012). Towards Operational Automatic Flood Detection Using EOS/MODIS Data. *Photogrammetric Engineering and Remote Sensing*, 78 to appear.
- Sun, D., Zhang, R., Li, S., & Yu, Y. (2011). GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Flood/Standing Water. Available at: http://www.goes-r.gov/products/ATBDs/option2/Land_FSW_v1.0_no_color.pdf (Accessed: August, 1st, 2011).
- USGS (2010). MODIS Reprojection Tool - Swath | Tools | LP DAAC :: ASTER and MODIS Land Data Products and Services. Available at: https://lpdaac.usgs.gov/lpdaac/tools/modis_reprojection_tool_swath (Accessed: Nov. 1, 2010).
- USGS (2011). USGS WaterWatch. Available at: <http://waterwatch.usgs.gov/new/> (Accessed: September 20, 2011).
- Vapnik, V. (2000). *The nature of statistical learning theory* (2nd ed.). New York: Springer.
- Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., et al. (2011). Validation of GOES-R Satellite Land Surface Temperature Algorithm Using SURFRAD Ground Measurements and Statistical Estimates of Error Properties. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 704–713.
- Zhang, R., & Ma, J. (2008). An improved SVM method P-SVM for classification of remotely sensed data. *International Journal of Remote Sensing*, 29, 6029–6036.
- Zhu, G., & Blumberg, D. G. (2002). Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel. *Remote Sensing of Environment*, 80, 233–240.